PHYSICAL REVIEW D **95,** 036012 (2017)

Simulating *b*-associated production of *Z* and Higgs bosons with the SHERPA event generator

Frank Krauss,[1] Davide Napoletano,[1] and Steffen Schumann[2]

[1]*Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, United Kingdom*
[2]*II. Physikalisches Institut, Georg-August-Universität Göttingen, 37077 Göttingen, Germany*

(Received 21 December 2016; published 16 February 2017)

We compare four- and five-flavor scheme predictions for *b*-associated production of *Z* and Higgs bosons. The results are obtained with SHERPA 's MC@NLO implementation for the four-flavor scheme, treating the *b*'s as massive, and with multijet merging at leading and next-to-leading order for the five-flavor schemes. Comparison with data for $Z + b(\bar{b})$ production at the 7 TeV LHC exhibits strengths and weaknesses of the different approaches and is used to validate predictions for *b*-associated Higgs-boson production at the 13 TeV Run II.

DOI: 10.1103/PhysRevD.95.036012

I. INTRODUCTION

The production of Z and H bosons in association with b quarks or through $b\bar{b}$-annihilation has recently attracted renewed interest, for a number of old and new reasons [1–4].

First, the associated production of a vector boson (V) and a b-tagged jet points to underlying processes like $gb \to Vb$ at Born level and is thus sensitive to the b-quark parton distribution function (PDF). These are particularly important for phenomenologically relevant processes as the production of a Higgs boson in $b\bar{b}$-annihilation, in association with b jets or a single top. The latter processes contribute to the total Higgs-boson production in the Standard Model on the level of a few percent and must therefore be included in fits to the couplings of the Higgs boson [3].

Second, and complementary to this purely Standard Model reasoning, many models for physics beyond the Standard Model come with extended Higgs sectors, quite often in the form of a second Higgs doublet. The mixing among the Higgs doublets often amplifies the couplings of the Higgs bosons to the b quarks. As a consequence, bH and $b\bar{b}H$ production provide important search grounds for new physics. Furthermore, events with identified b jets and a significant missing transverse momentum constitute a possible signature of dark matter [5–7] production. For this signal, invisibly decaying Z bosons associated with b jets pose a severe irreducible background.

In addition to the processes considered in this work, also hadronic single top-quark production through the t-channel process proceeds via initial-state b quarks. At leading order, this corresponds to the process $qb \to tq'$. A proper treatment of the initial-state b quark and of higher-order processes such as $qg \to tq'\bar{b}$ is again vital for the successful description of this important signature.

In all cases, however, the treatment of the b quark is far from being straightforward, since commonly PDFs are assumed to be valid for massless partons only—parton masses induce logarithmic and unknown power corrections. On the other hand, the b-quark mass is large enough, around 4.5 GeV, to induce visible kinematic effects for jets. With a jet transverse momentum of about 20 GeV, they can be estimated by $(m_b/p_T)^2$, on the level of around 10%. Aiming for such accuracies, the treatment of the b-quark mass therefore poses a problem. One solution is the five-flavor scheme [8], defined by assuming the b quark as strictly massless in the matrix elements and by allowing a nonvanishing b-quark PDF. In the context of our studies, this translates into using multijet-merging technology to combine processes such as $b\bar{b} \to H$, $gb \to Hb$, $gg \to Hb\bar{b}$, etc., with $m_b = 0$ into a fully inclusive sample. On the other hand, from an alternative point of view, one could also claim that b-quark PDFs were ill-defined objects and could therefore set them to zero. In this case, one would study processes such as $gg \to Zb\bar{b}$ and $gg \to Hb\bar{b}$ [9,10] instead and use the finite b-quark mass to regularize the otherwise divergent phase-space integrals. This treatment, namely taking the b-quark mass fully into account but releasing any phase-space constraints on their final state, defines the four-flavor scheme.

Various ways of combining results obtained in these two schemes have been proposed. Among them, the so-called *Santander matching* first presented in Ref. [11] is probably the most widely used one. Its approach is to combine four- and five-flavor scheme predictions by means of a dynamically weighted average of them. This weight is defined to be a continuous function of the hard scale of the process and the mass of the bottom quark, in such a way that, when the ratio between the hard scale and the b-quark mass is large, the five-flavor (5F) prediction is recovered, while for scales of similar size, the four-flavor (4F) result is obtained.

As a well-defined alternative, the FONLL approach has been introduced for b-quark hadroproduction [12]. It was later extended to deep-inelastic scattering (DIS) [13],

and lately it has been applied to Higgs production in bottom-quark fusion [14,15]. The main idea of this method is to take the four- and five-flavor scheme perturbative-series expansions and, after having rearranged them in such a way that they become compatible, to isolate double-counting terms in the two schemes. This is achieved by reexpressing PDF evolution and the running of α_S in the four-flavor scheme in terms of those computed in the five-flavor scheme and heavy-flavor PDFs in the five-flavor scheme in terms of light-flavor PDFs. The final prediction is obtained by replacing terms in the five-flavor scheme by their known counterparts computed in the four-flavor scheme.

Another matching procedure, based on an Effective Field Theory approach, has recently been developed in Refs. [16,17]. More methods to match initial-state resummed massless predictions with fixed-order ones are available in DIS physics. In general, decoupling schemes, like the four-flavor scheme, in this context, are referred to as fixed-flavor-number schemes, while massless schemes, in which the number of active flavors changes with energy, like the five-flavor scheme, are called variable-flavor-number schemes (VFNS). Adding mass effects to a VFNS leads to a general-mass (GM) VFNS (as opposed to the completely massless zero-mass VFNS. Both Santander and the FONLL matching together with the ACOT scheme in different versions [18–20] fall into the GM VFNS class.

This discussion and the corresponding schemes apply mainly to analytic calculations. As soon as fragmentation effects are to be accounted for, the finite b-quark mass must be taken into account, as it reduces the emission rate of gluons off the quark with respect to the strictly massless case. To avoid a resulting fragmentation function which would be significantly too soft, b quarks need to be treated as massive within parton-shower simulations. Contact with massless calculations is established by shifts of the 4-momenta of b quarks before or during their first emission.

The simulation of final states with variable jet multiplicities, i.e. varying levels of inclusiveness, is the realm of matrix-element parton-shower matching and merging techniques [21]. They combine the strengths of both approaches. Exact leading-order (LO) or next-to-leading-order (NLO) QCD matrix elements describe hard, well-separated parton configurations, while additional softer jets and in general jet evolution are accounted for by parton showers. Considering final states with *identified* b jets certainly poses stringent tests on these algorithms and in fact requires dedicated methods to correctly account for the nonvanishing b-quark mass in both ingredients of the calculations.

In this publication, we discuss and validate the corresponding methods within the SHERPA event generator [22,23]. Different choices for treating the b quarks in the matrix elements, ranging from massless in a five-flavor scheme (5FS) to massive in a four-flavor scheme (4FS), consistently combined with parton showers, are compared. The presented approaches are implemented and readily available from SHERPA-2.2.1.

The discussion is organized as follows. In Sec. II, the underlying calculations are briefly reviewed. In addition, the different flavor schemes are defined in more detail. In Sec. III, using $Z + b$ and $Z + b\bar{b}$ data at 7 TeV from Run I of the LHC, the methods are validated, and their relative strengths and weaknesses are identified. The findings will be used in the next section, Sec. IV, to arrive at informed and robust predictions for the b-associated production of Higgs bosons at the 13 TeV Run II of the LHC. In the Conclusions, some comments concerning further implications for beyond the Standard Model physics put this study into a wider context.

II. CALCULATIONAL METHODS AND SETUPS

Efficient routines for the required QCD matrix-element calculations and a well-understood QCD parton shower are the key ingredients to all matching and merging calculations. Within SHERPA, LO matrix elements are provided by the built-in generators AMEGIC++ [24] and COMIX [25]. While virtual matrix elements contributing to QCD NLO corrections can be invoked through interfaces to a number of specialized tools, e.g. BLACKHAT [26], GOSAM [27], NJET [28], OPENLOOPS [29], or through the BLHA interface [30], we employ in this study the OPENLOOPS generator [31] in conjunction with the COLLIER library [32,33]. Infrared divergences are treated by the Catani-Seymour dipole method [34,35], which has been automated in SHERPA [36]. In this implementation, mass effects are included for final-state splitter and spectator partons, but massless initial-state particles are assumed throughout. SHERPA's default parton-shower model [37,38] is based on Catani-Seymour factorization [39]. In order to arrive at meaningful fragmentation functions for heavy quarks, all modern parton showers take full account of their finite masses in the final state, although in algorithmically different ways. In SHERPA, the transition from massless to massive kinematics is achieved by rescaling 4-momenta at the beginning of the parton shower. In the initial-state parton shower in SHERPA, the $g \to b\bar{b}$ and $b \to bg$ splitting functions do not contain b-quark mass effects in their functional form and account for mass effects in the kinematics only.

In the following, we briefly define the methods available in SHERPA for simulating b-associated production processes, that will then be validated and applied for LHC predictions:

(i) *4F NLO (4F MC@NLO).*—In the *four-flavor scheme*, b quarks are consistently treated as *massive* particles, only appearing in the final state. As a consequence, b-associated Z- and H-boson production proceeds through the parton-level processes

$gg \to Z/H + b\bar{b}$, and $q\bar{q} \to Z/H + b\bar{b}$ at Born level. MC@NLO matching is obtained by consistently combining fully differential NLO QCD calculations with the parton shower, cf. Refs. [40,41]. Due to the finite b-quark mass, these processes do not exhibit infrared divergences, and the corresponding *inclusive* cross sections can thus be evaluated without any cuts on the b-partons.

(ii) *5F LO (5F MEPS@LO)*: In the *five-flavor scheme*, b quarks are *massless* particles in the *hard matrix element*, while they are treated as massive particles in both the initial- and final-state *parton shower*. In the MEPS@LO [42] samples, we merge $pp \to H/Z$ plus up to three jets at leading order; this includes, for instance, the parton-level processes $b\bar{b} \to Z/H$, $gb \to Z/Hb$, $gg \to Z/Hb\bar{b}, \ldots$. To separate the various matrix-element multiplicities, independent of the jet flavor, a jet cut of $Q_{\text{cut}} = 10$ GeV is used in the Z case while $Q_{\text{cut}} = 20$ GeV is employed in H-boson production.

(iii) *5F NLO (5F MEPS@NLO)*: In the 5FS MEPS@NLO scheme [43,44], we account for quark masses in complete analogy to the LO case: the quarks are treated as massless in the hard matrix elements but as massive in the initial- and final-state parton showering. Again, partonic processes of different multiplicity are merged similarly to the MEPS@LO albeit retaining their next-to-leading-order accuracy. In particular, we consider the merging of the processes $pp \to H/Z$ plus up to two jets each calculated with MC@NLO accuracy further merged with $pp \to H/Z + 3j$ calculated at MEPS@LO.

We consistently use four-flavor PDFs in the 4FS, i.e. the dedicated four-flavor NNPDF3.0 set [45] with the strong coupling given by $\alpha_s(m_Z) = 0.118$ and running at NLO. For the simulations in the five-flavor schemes, the five-flavor next-to-next-to-leading-order (NNLO) PDFs from NNPDF3.0 are used, with $\alpha_s(m_Z) = 0.118$ and running at NNLO. We assume all quarks apart from the b to be massless, with a pole mass of $m_b = 4.92$ GeV which enters the hard matrix-element calculation, where appropriate, and the parton shower.

Results in the 4F and 5F schemes have been obtained with the default scale-setting prescription for parton-shower matched calculations in SHERPA [42,46]. They are calculated using a backward-clustering algorithm, and for each emission from the shower, couplings are evaluated at either the k_T of the corresponding emitted particle (in the case of gluon emission) or at the invariant mass of the emitted pair (in the case of gluon splitting into quarks). The clustering stops at a "core" $2 \to 2$ process, with all scales set to $\mu_F = \mu_R = \mu_Q = m_T(V)/2$, where $m_T(V)$ corresponds to the transverse mass of the boson. This scale is thus used to evaluate couplings in the hard matrix element and PDFs. The corresponding central values are supplemented with uncertainty bands reflecting the dependence on the unphysical scales. Renormalization and factorization scales are varied around their central value by a factor of 2 up and down, with a standard seven-point variation. The scale variations use the SHERPA internal reweighting procedure [47] and result in envelopes around the central value. Furthermore, we consider explicit variations of the parton-shower starting scale, i.e. μ_Q, by a factor of 2 up and down.

For the Higgs-bosons production processes, bottom and top Yukawa couplings are important. Using their corresponding pole mass, $m_b = 4.92$ GeV and $m_t = 172.5$ GeV, and subject to a LO running, they are finally evaluated at $\mu_m = m_H = 125$ GeV. We do not include variations of this scale.

III. BOTTOM-JET ASSOCIATED Z-BOSON PRODUCTION

The production of a Z boson in association with QCD jets provides the ideal test bed for the theoretical approaches outlined above. Through the decay of the Z boson to leptons, these processes yield a rather simple and clean signature with sizeable rates even for higher jet counts. Precise measurements of the production rates and differential distributions of both the Z-boson decay products and the accompanying jets offer discriminating power for miscellaneous theoretical approaches. In fact, measurements of Z + jets production served as key inputs for the validation of matrix-element parton-shower simulation techniques, cf. Refs. [48–50], and impressively underpin the enormous success of these calculational methods.

Here, we focus on the production of Z bosons accompanied by identified b jets. Comparison with data from both the ATLAS and CMS collaborations at 7 TeV [51,52] provides the benchmark for the accuracy and quality of four- and five-flavor simulations with SHERPA. Similar measurements at 8 and 13 TeV LHC collision energies are under way [53].

A. Measurements at LHC Run I—Reference data

Based on a data set of 4.6 fb^{-1} integrated luminosity. the ATLAS Collaboration studied the production of b jets associated with Z/γ^* that decay to electrons or muons [51]. The dilepton invariant mass in the range 76 GeV $< m_{\ell\ell} <$ 106 GeV. Jets are reconstructed using the anti-k_t algorithm [54] with a radius parameter of $R = 0.4$, a minimal transverse momentum of $p_{T,j} > 20$ GeV, and a rapidity of $|y_j| < 2.4$. Furthermore, each jet candidate needs to be separated from the leptons by $\Delta R_{j\ell} > 0.5$. Jets containing b-hadrons are identified using a multivariate technique. To match the outcome of the experimental analysis, simulated jets are identified as b jets, when there is one or more weakly decaying b-hadron with $p_T >$ 5 GeV within a cone of $\Delta R = 0.3$ around the jet axis.

The sample of selected events is further subdivided into a class containing events with at least one b jet (1-tag) and a class with at least two b jets (2-tag).

A similar analysis was performed by CMS [52]. There, electrons and muons are required to have a transverse momentum of $p_{T,\ell} > 20$ GeV, a pseudorapidity $|\eta_\ell| < 2.4$, and a dilepton invariant mass within 81 GeV $< m_{\ell\ell} <$ 101 GeV. Only events with exactly two additional b-hadrons were selected. The analysis focuses on the measurement of angular correlations among the b-hadrons and with respect to the Z boson. This includes in particular variables sensitive to rather collinear b-hadron pairs. In addition, the total production cross section as a function of the vector boson's transverse momentum was measured.

Both analyses are implemented and publicly available in the RIVET analysis software [55] that, together with the FASTJET package [56], is employed for all particle-, i.e. hadron-, level analyses in this publication.

B. Comparison with LHC data

In this section, the theoretical predictions from SHERPA will be compared to the experimental measurements from LHC Run I. We begin the discussion with the comparison with the measurements presented by the ATLAS Collaboration in Ref. [51]. The total cross sections for $Z+ \geq 1$ and $Z+ \geq 2$ b jets are collected in Fig. 1. Already, there, we see a pattern emerging that will further establish itself in the differential cross sections: while the five-flavor MEPS@NLO results agree very well with the data, the central values of the five-flavor MEPS@LO cross sections tend to be around 10%–20% lower than the central values of the data, but with theory uncertainties clearly overlapping them. For all the runs, the uncertainty estimates include both seven-point variations of the perturbative scales $\mu_{R/F}$ as well as μ_Q variations by a factor of 2 up and down. In contrast to the five-flavor case, the four-flavor MC@NLO cross sections tend to be significantly below the experimental values for the $Z+ \geq 1$ b-jets cross section, without overlap of uncertainties. In the $Z+ \geq 2$ b-jets cross section, the agreement between four-flavor MC@NLO results and the data is better, with the theoretical uncertainties including the central value of the measured cross section.

In Fig. 2, the differential cross sections with respect to the transverse momentum and rapidity of the b jets, normalized to the number of b jets, are presented for events with *at least* one b-tagged jet. The shapes of both distributions are well modelled both by the 4F and the two 5F calculations. However, clear differences in the predicted production cross sections are observed. While the 5F NLO results are in excellent agreement with data—both in shape and normalization—the central values of the 5F LO cross sections tend to be around 10% below data, at the lower edge of the data uncertainty bands, and the 4F results are consistently outside the data, about 25% too low. In the lower panels of Fig. 2 and all the following plots in this section, we show the uncertainty bands of the theoretical predictions, corresponding to the above-described $\mu_{R/F}$ and μ_Q variations. For the 5FS calculations, the scale uncertainties clearly dominate, while for the 4F MC@NLO scheme, the shower-resummation uncertainty dominates.

This pattern is repeated in Fig. 3, where we show the differential $\sigma(Zb)$ cross section with respect to the dilepton transverse momentum and, rescaled to $1/N_{b-\text{jets}}$, as a

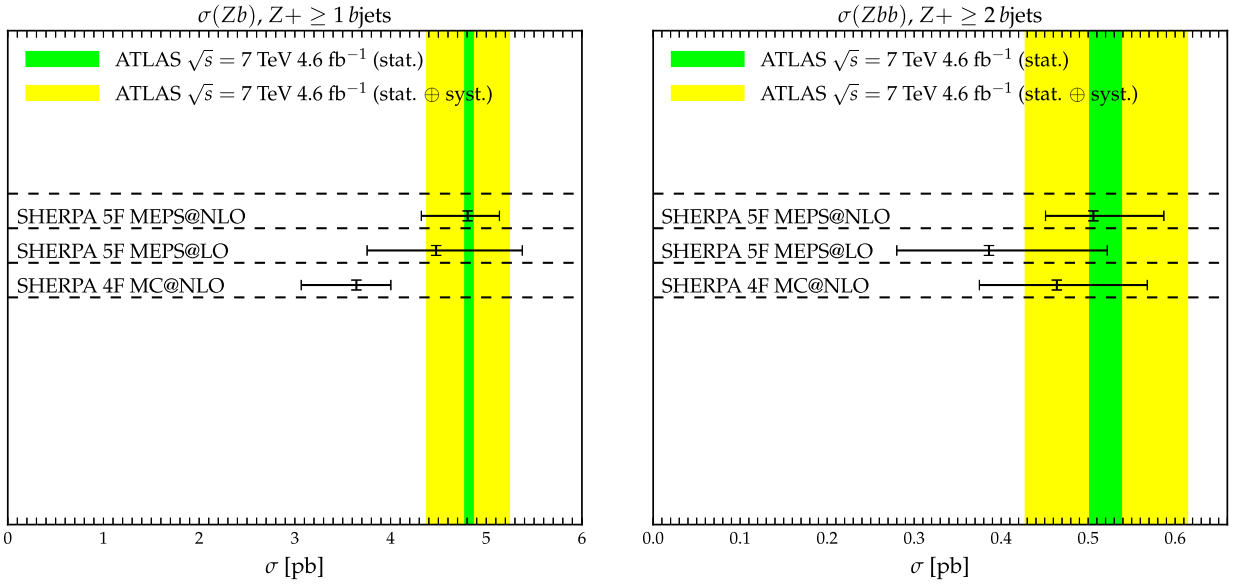

FIG. 1. Comparison of total production cross section predictions with ATLAS data [51]. The error bars on the theoretical results are calculated from variations of the hard-process scales $\mu_{R/F}$ and the parton-shower starting scale μ_Q.

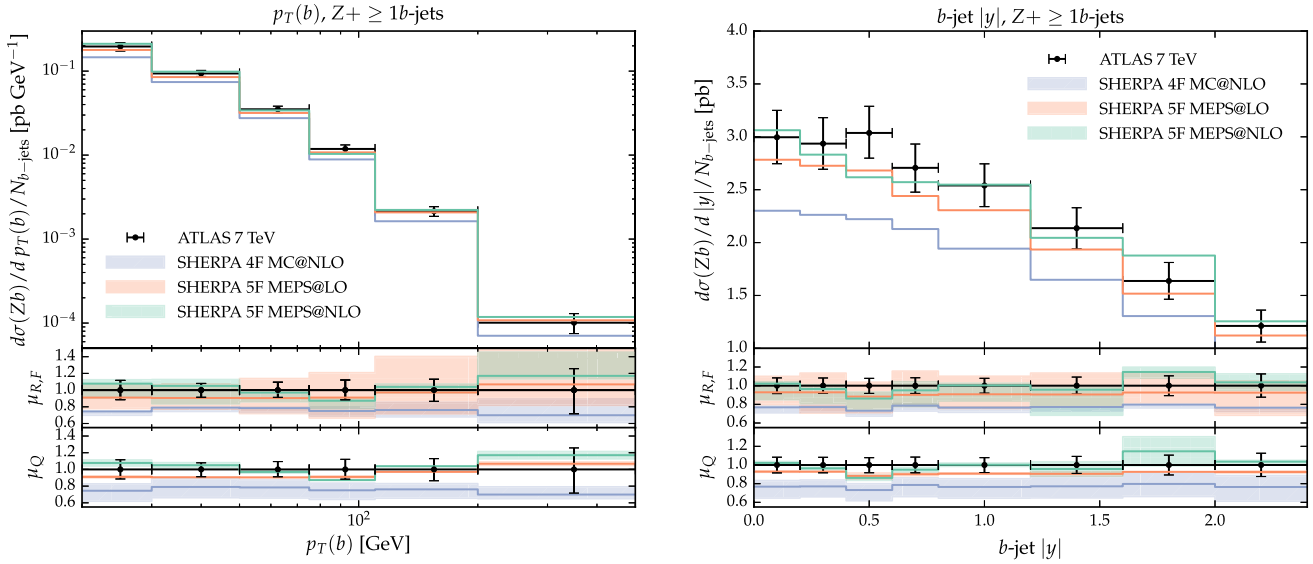

FIG. 2. Inclusive transverse-momentum and rapidity distribution of all b jets in events with at least one b jet. Data are taken from Ref. [51].

function of the azimuthal separation between the reconstructed Z boson and the b jets. Again, both distributions are very well modelled by both 5F calculations. The 4F MC@NLO prediction again underestimates data by a largely flat 20%–25%.

Moving on to final states exhibiting at least two identified b jets, the role of the 5F LO and 4F NLO predictions are somewhat reversed: as can be inferred from Fig. 1, the 4F and 5F NLO samples provide good estimates for the inclusive Zbb cross section, while the 5F LO calculation undershoots the data by about 20%. In Fig. 4, the ΔR separation of the two highest transverse-momentum b jets along with their invariant-mass distribution is presented. Both the 4F and the 5F approaches yield a good description of the shape of the distributions. It is worth stressing that this includes the regions of low invariant mass and low ΔR, corresponding to a pair of rather collinear b jets. This is a region that is usually riddled by potentially large logarithms, where the parton shower starts taking effect. Note that in the comparison presented

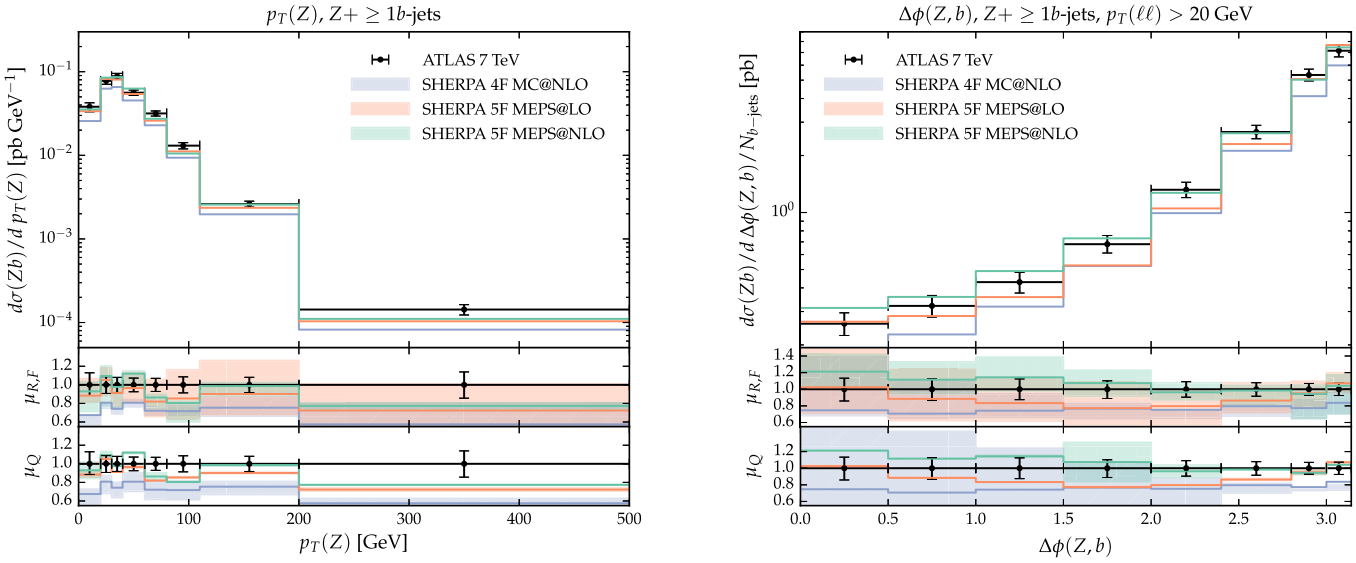

FIG. 3. Transverse-momentum distribution of the Z boson (left) and the azimuthal separation between the Z boson and the b jets (right) in events with at least one b jet. For the $\Delta\phi(Z,b)$ measurement, the additional constraint $p_{T,ll} > 20$ GeV is imposed. Data are taken from Ref. [51].

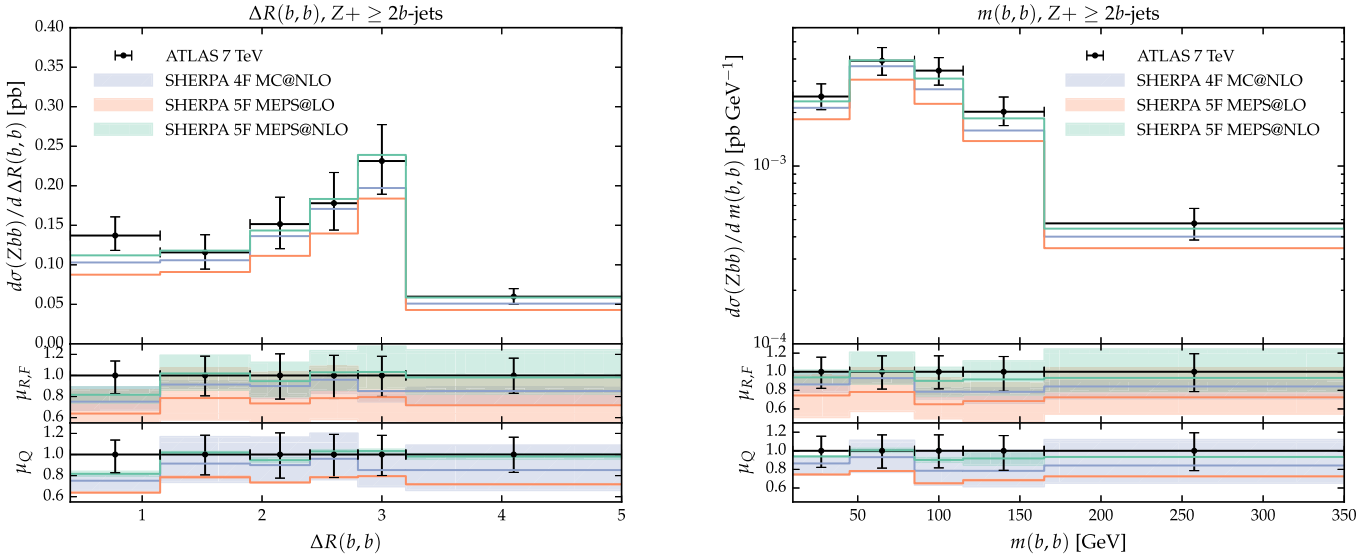

FIG. 4. The ΔR separation (left) and invariant-mass distribution (right) for the leading two b jets. Data are taken from Ref. [51].

in Ref. [51] this region showed some disagreement between the data and other theoretical predictions based on NLO QCD (dressed with parton showers).

In Fig. 5, the resulting transverse-momentum distribution of the dilepton system when selecting for events with at least two associated b jets is shown. The shape of the data is very well reproduced by the 4F MC@NLO and 5F MEPS@NLO samples. Also, the 5F MEPS@LO prediction describes the data well, despite the overall rate being 20% lower than what is observed in the data.

The measurements presented by the CMS Collaboration in Ref. [52] focus on angular correlations between b-hadrons rather than b jets. Two selections with respect to the dilepton transverse momentum have been considered, a sample requiring $p_T(Z) > 50$ GeV and an inclusive one considering the whole range of $p_T(Z)$. The ΔR and $\Delta\phi$ separation of the b-hadrons obviously proves to be most sensitive to the theoretical modelling of the b-hadron production mechanism and the interplay of the fixed-order components and the parton showers. They are presented in Figs. 6 and 7. In general, a good agreement in the shapes of simulation results and data is found, with the same pattern of total cross sections as before: the 5F MEPS@NLO sample describes the data very well, while the 4F MC@NLO results tend to be a little bit, about 10%, below data, with data and theory uncertainty bands well overlapping, while the central values of the 5F MEPS@LO results undershoot the data by typically 20%–25%.

Overall, it can be concluded that the 5F MEPS@NLO calculation yields the best description of the existing measurements, regarding both the production rates *and* shapes. The 4F MC@NLO and 5F MEPS@LO schemes successfully model the shape of the differential distributions but consistently underestimate the production rates.

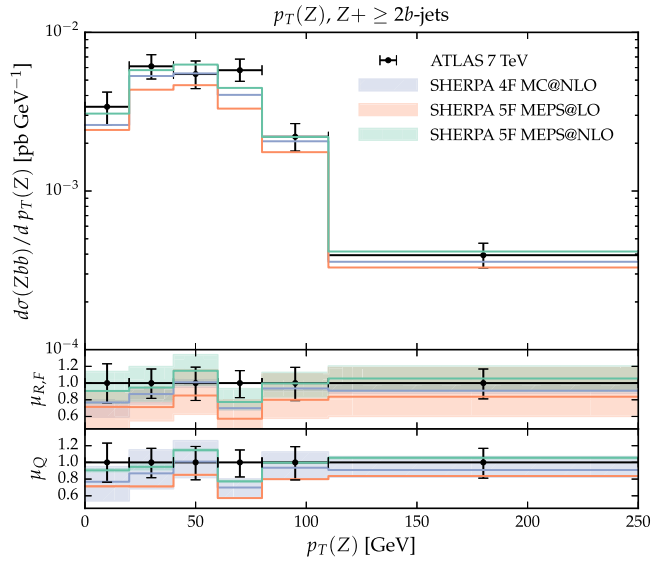

FIG. 5. Transverse-momentum distribution of the dilepton system for events with at least two b jets. Comparison is against various calculational schemes. Data are taken from Ref. [51].

IV. BOTTOM-JET ASSOCIATED HIGGS-BOSON PRODUCTION

In this section, we present predictions for b-jet(s) associated production of the Standard Model Higgs boson in pp collisions at the 13 TeV LHC obtained in the four- and five-flavor schemes. As is standard when dealing with this process, we do not include contributions from the gluon-fusion channel. However, in the 4F MC@NLO, we do include terms proportional to the top-quark Yukawa coupling, contributing to order $y_b y_t$ as an interference effect at NLO QCD [3,57,58]. Although associated $Z + b$-jet(s) production serves as a good proxy for the Higgs-boson

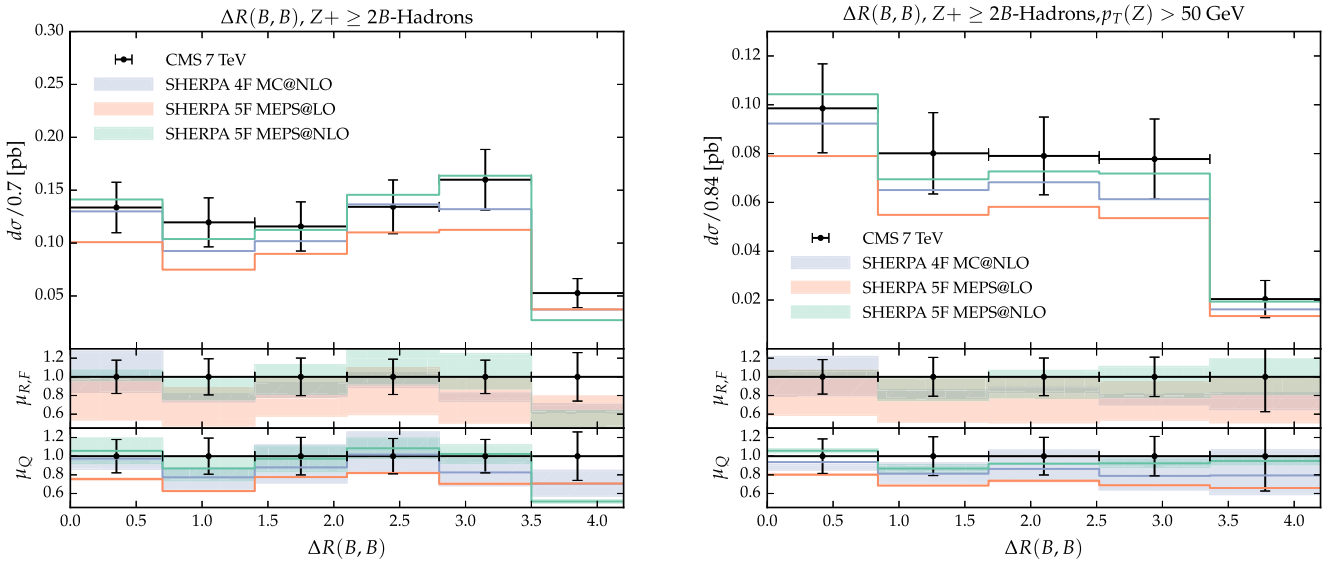

FIG. 6. ΔR_{BB} distribution for two selections of the transverse momentum of the Z boson. Data are taken from Ref. [52].

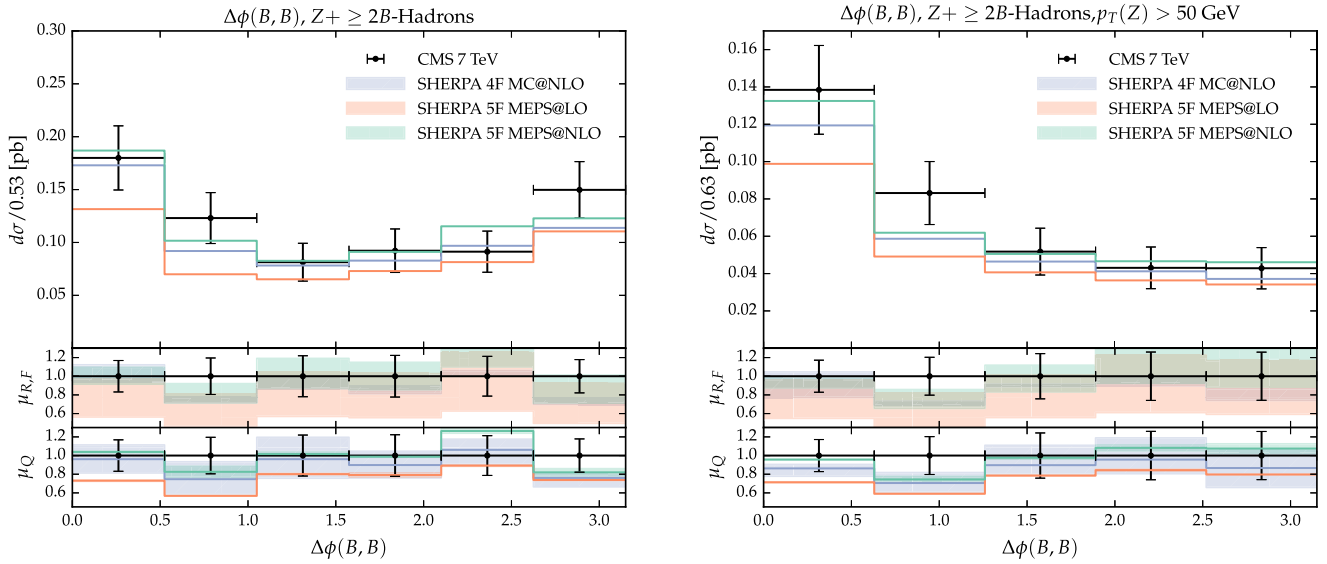

FIG. 7. $\Delta\phi_{BB}$ distribution for two selections of the transverse momentum of the Z boson. Data are taken from Ref. [52].

case, there are important differences between both processes, mainly due to the different impact of initial-state light quarks, which couple to Z bosons but not to the Higgs boson.

As before, QCD jets are defined through the anti-k_t algorithm using a radius parameter of $R = 0.4$, a minimal transverse momentum $p_{T,j} > 25$ GeV, and a rapidity cut of $|y_j| < 2.5$. In this case, we consider results that are at the parton level only, disregarding hadronization and underlying-event effects, which may blur the picture. We consider single b-tagged jets only, thus excluding jets with intrajet $g \to b\bar{b}$ splittings from the parton shower which would be the same for all flavor schemes we investigate. As for Z-boson production, we separate the event samples into categories with at least one b jet, i.e. $H+ \geq 1b$-jet events, and at least two tagged b jets, i.e. $H+ \geq 2b$-jets events.

In Table I, cross sections for the three calculations are reported. Historically, inclusive results have largely disagreed between the 4F and the 5F scheme. This feature is observed for the case at hand, too, and especially so for the case of one tagged b jet. There, the 4F MC@NLO prediction is smaller than the 5F results by factors of about 1.75 (5F LO) and of 2.44 (5F NLO). The relative differences are reduced when a second tagged b jet is demanded. In this case, we find that the 4F result lies between the two 5F results, about 20% higher than the LO predictions, and a

TABLE I. 13 TeV total cross sections and the corresponding $\mu_{F/R}$ and μ_Q uncertainties for $H+ \geq 1b$ and $H+ \geq 2b$s.

LHC 13 TeV	$H+ \geq 1b$ jets (fb)	$H+ \geq 2b$ jets (fb)
$\sigma^{4F}_{\text{MC@NLO}}$	$45.2^{+15.5\%}_{-18.4\%}$	$4.5^{+25.1\%}_{-26.3\%}$
$\sigma^{5F}_{\text{MEPS@LO}}$	$79.3^{+34.0\%}_{-25.4\%}$	$3.8^{+34.3\%}_{-30.3\%}$
$\sigma^{5F}_{\text{MEPS@NLO}}$	$110.5^{+14.2\%}_{-16.0\%}$	$6.9^{+27.3\%}_{-27.1\%}$

factor of about 1.5 lower than the 5F NLO predictions. In both cases, inclusive $H + b$ and $H + bb$ production, the uncertainty bands of the two 5F predictions, corresponding to seven-point $\mu_{R/F}$ variations and μ_Q variations by a factor of 2 up and down, do overlap. While for the two b-jet final states this includes the 4F result, for the one b-jet case, the 4F result is not compatible with the 5F predictions, taking into account the considered scale uncertainties. It is worth noting that a milder form of this relative scaling of the cross sections was already observed in the Z case.

In the case of the total inclusive cross section, this very large difference can be mitigated by including higher-order corrections, on the one hand, and a better assessment of which choice of the unphysical scales yields the better agreement [1–4]. However, only a recent effort to match the two schemes [14–17] has clearly assessed the relative importance of mass corrections (appearing in the 4F scheme) and large log resummation (as achieved in a 5F scheme). In particular, it has been found that the difference between these two schemes is mostly given by the resummation of large logarithms, thus suggesting that for an inclusive enough calculation either a 5F scheme or a matched scheme should be employed. This is the same situation that one faces, albeit milder, in the Z case, where, in terms of normalization, the 5F scheme performed better in all cases and especially in inclusive calculations. We therefore recommend that in terms of overall normalization the 5F MEPS@NLO scheme should be used to obtain reliable predictions.

Let us now turn to the discussion of the relative differences in the shapes of characteristic and important distributions. To better appreciate shape differences, all differential distribution are normalized to the respective cross section, i.e. the inclusive rates $\sigma(Hb)$ and $\sigma(Hbb)$. In all cases, we obtain agreement at the 15% level or better between the 5F MEPS@NLO and 4F MC@NLO samples, the only exception, not surprisingly, being the region of phase space where the two b's come close to each other and resummation effects start playing a role. Typically, the 5F MEPS@LO predictions are also in fair agreement with the other two results; however, they exhibit a tendency for harder tails in the p_T distributions, mainly in the inclusive Higgs-boson p_T and in the transverse momentum of the second b jet.

Starting with Fig. 8, the transverse-momentum distributions of the Higgs boson and the leading b jet in the case of at least one b jet tagged are displayed. Similarly to the Z example, this is the region where one would expect the 5F scheme to perform better. However, again similarly to the Z case, the three schemes largely agree in terms of shapes, being well within scale uncertainties. Notably, this turns out to be particularly true for the low (∼20–100 GeV) p_T region where one could have expected deviations to be the largest.

In Figs. 9 and 10, we present differential distributions for the selection of events with at least two tagged b jets. While Fig. 9 shows the resulting Higgs-boson transverse-momentum distribution, Fig. 10 compiles results for the ΔR separation of the two leading b jets and their invariant-mass distribution. For two such b-jets observables, the 4F scheme

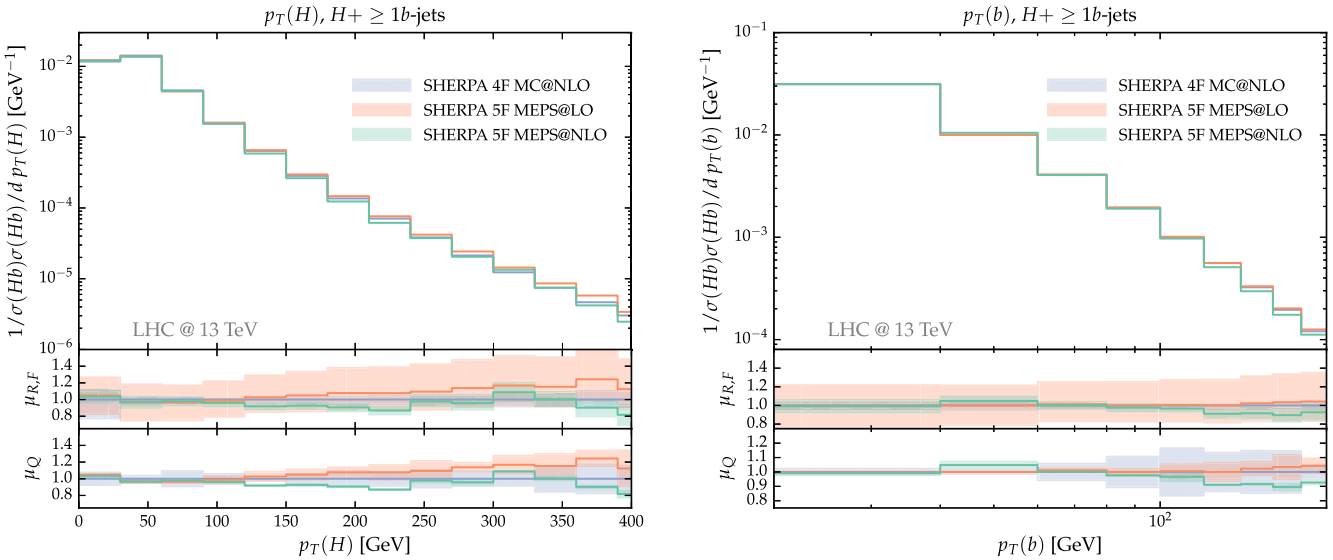

FIG. 8. Predictions for the transverse-momentum distribution of the Higgs boson (left panel) and the leading b jet (right panel) in inclusive $H + b$-jet production at the 13 TeV LHC.

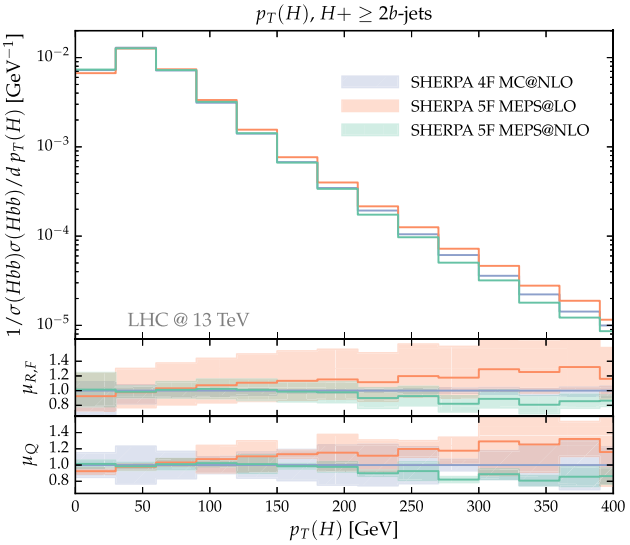

FIG. 9. The transverse-momentum distribution of the Higgs boson in inclusive $H + 2b$-jets production at the 13 TeV LHC.

is expected to work best, especially when the two b are well separated to suppress potentially large logarithms. However, in agreement with the Z-boson case, no significant differences between the various schemes arise when taking into account $\mu_{R/F}$ and μ_Q scale-variation uncertainties. Once again, the region of low p_T in Fig. 9 and the region of low $m(b,b)$ in Fig. 10 show excellent agreement among the various descriptions. As anticipated, larger differences can be seen between the two 5FS and the 4F MC@NLO calculations, in the very low $\Delta R(b,b)$ and $m(b,b)$ regions, Fig. 10, where the two b jets become collinear. This feature is, however, most likely due to the fact that we are dealing with partonic b jets as opposed to *hadronic* ones. Taking as a reference the Z-boson case once again, in fact, where this difference is not present at all, suggests that a realistic simulation, that accounts for hadronization effects, should largely suppress this difference.

V. CONCLUSIONS

Simulations for the associated production of a Z or a Higgs boson with a $b\bar{b}$ pair have always proven to require careful thinking in including or neglecting b-quark mass effects. In this work, a detailed comparison between the 4F and the 5F schemes implemented in SHERPA has been presented.

Firstly, the results for production of a Z boson with b jets has been compared with both ATLAS and CMS data. We find that all schemes largely agree in the shapes of relevant observables. Major differences, however, appear in the overall normalization of the various samples, with the 5F MEPS@NLO prediction being the one proving the best agreement with data.

We used this as a guide to study the b-associated production of Standard Model Higgs bosons. Due to the different impacts of the initial-state b quarks, this process enhances the quantitative differences between the different approaches, and in particular the production cross sections, while it still maintains the qualitative scaling behavior. This qualitative similarity is fortified by the good agreement of the calculations in the shapes of sensitive observables.

We thus conclude that, in order to obtain reliable predictions, at the LHC, for the production of a Higgs boson with b jets, the use of a 5F MEPS@NLO setup is the most advisable. A second, more phenomenologically driven, option could be to use a 4FS, MC@NLO accurate, prediction, normalized by the 5F MEPS@NLO total cross

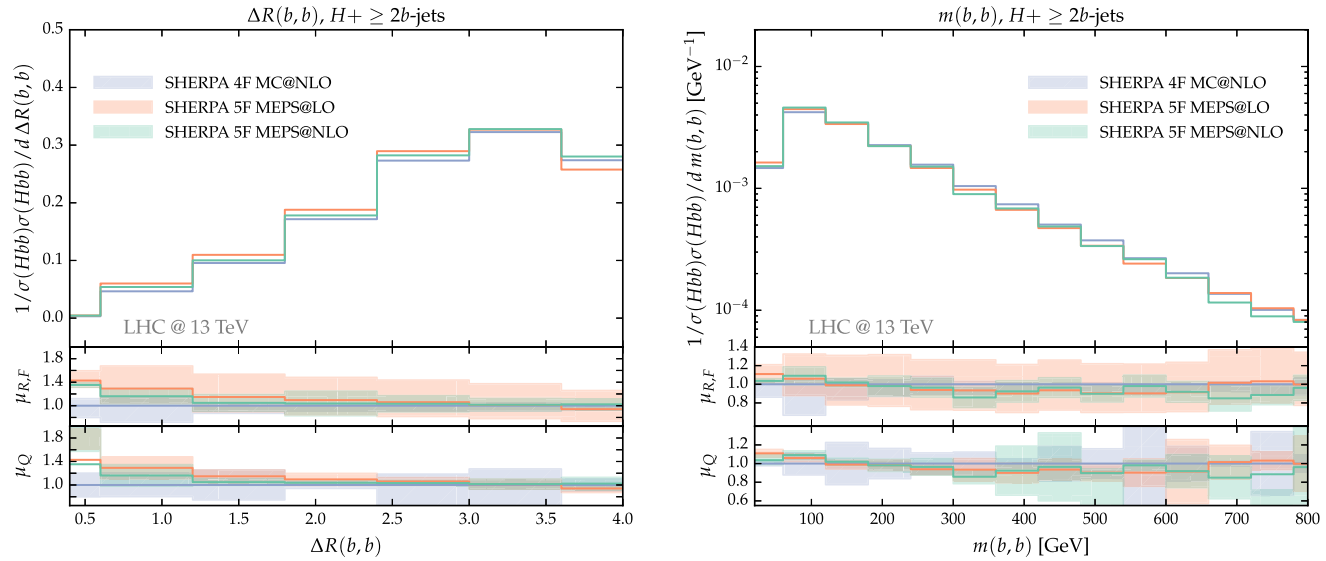

FIG. 10. Predictions for the ΔR separation of the two leading b jets (left panel) and their invariant-mass distribution (right panel) in inclusive $H + 2b$-jets production at the 13 TeV LHC.

section. This is particularly relevant given that the 4FS calculation is by far the most efficient one.

To further improve our theoretical predictions, we plan to generalize the treatment of finite-mass effects in the 5FS by allowing for massive initial-state quarks in the matrix elements and the parton showers. Besides using fully mass-dependent matrix elements also at NLO, this requires a generalization of the implementation of the NLO subtraction formalism, along the lines of Ref. [59], as well as the inclusion of mass effects in the initial-state parton-shower splitting functions.

ACKNOWLEDGMENTS

We want to thank our colleagues from the SHERPA Collaboration for fruitful discussions and technical support. We acknowledge financial support from the EU research networks funded by the Research Executive Agency of the European Union under Grants No. PITN-GA2012-316704 ("HiggsTools") and No. PITN-GA-2012-315877 ("MCnetITN"), by the ERC Advanced Grant MC@NNLO (Grant No. 340983), and from BMBF under Contracts No. 05H12MG5 and No. 05H15MGCAA.

[1] R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, R. Pittau, and P. Torrielli, W and $Z/\gamma*$ boson production in association with a bottom-antibottom pair, J. High Energy Phys. 09 (2011) 061.

[2] M. Wiesemann, R. Frederix, S. Frixione, V. Hirschi, F. Maltoni, and P. Torrielli, Higgs production in association with bottom quarks, J. High Energy Phys. 02 (2015) 132.

[3] D. de Florian *et al.* (LHC Higgs Cross Section Working Group Collaboration), Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector, arXiv:1610.07922.

[4] M. Lim, F. Maltoni, G. Ridolfi, and M. Ubiali, Anatomy of double heavy-quark initiated processes, J. High Energy Phys. 09 (2016) 132.

[5] T. Lin, E. W. Kolb, and L.-T. Wang, Probing dark matter couplings to top and bottom quarks at the LHC, Phys. Rev. D **88**, 063510 (2013).

[6] G. Aad *et al.* (ATLAS Collaboration), Search for dark matter in events with heavy quarks and missing transverse momentum in pp collisions with the ATLAS detector, Eur. Phys. J. C **75**, 92 (2015).

[7] N. Chen, Z. Kang, and J. Li, Missing particle associated with two bottom quarks at the LHC: Mono-b versus $2b$ with razor variables, Phys. Rev. D **95**, 015003 (2017).

[8] R. V. Harlander and W. B. Kilgore, Higgs boson production in bottom quark fusion at next-to-next-to leading order, Phys. Rev. D **68**, 013001 (2003).

[9] S. Dawson, C. Jackson, L. Orr, L. Reina, and D. Wackeroth, Associated Higgs production with top quarks at the large hadron collider: NLO QCD corrections, Phys. Rev. D **68**, 034022 (2003).

[10] F. Febres Cordero, L. Reina, and D. Wackeroth, NLO QCD corrections to W boson production with a massive b-quark jet pair at the Tevatron $p\bar{p}$ collider, Phys. Rev. D **74**, 034007 (2006).

[11] R. Harlander, M. Krämer, and M. Schumacher, Bottom-quark associated Higgs-boson production: reconciling the four- and five-flavour scheme approach, arXiv:1112.3478.

[12] M. Cacciari, M. Greco, and P. Nason, The P(T) spectrum in heavy flavor hadroproduction, J. High Energy Phys. 05 (1998) 007.

[13] S. Forte, E. Laenen, P. Nason, and J. Rojo, Heavy quarks in deep-inelastic scattering, Nucl. Phys. **B834**, 116 (2010).

[14] S. Forte, D. Napoletano, and M. Ubiali, Higgs production in bottom-quark fusion in a matched scheme, Phys. Lett. B **751**, 331 (2015).

[15] S. Forte, D. Napoletano, and M. Ubiali, Higgs production in bottom-quark fusion: matching beyond leading order, Phys. Lett. B **763**, 190 (2016).

[16] M. Bonvini, A. S. Papanastasiou, and F. J. Tackmann, Resummation and matching of b-quark mass effects in $b\bar{b}H$ production, J. High Energy Phys. 11 (2015) 196.

[17] M. Bonvini, A. S. Papanastasiou, and F. J. Tackmann, Matched predictions for the $b\bar{b}H$ cross section at the 13 TeV LHC, J. High Energy Phys. 10 (2016) 053.

[18] M. A. G. Aivazis, J. C. Collins, F. I. Olness, and W.-K. Tung, Leptoproduction of heavy quarks. II. A unified QCD formulation of charged and neutral current processes from fixed-target to collider energies, Phys. Rev. D **50**, 3102 (1994).

[19] M. A. G. Aivazis, F. I. Olness, and W.-K. Tung, Leptoproduction of heavy quarks. I. General formalism and kinematics of charged current and neutral current production processes, Phys. Rev. D **50**, 3085 (1994).

[20] M. Krämer, F. I. Olness, and D. E. Soper, Treatment of heavy quarks in deeply inelastic scattering, Phys. Rev. D **62**, 096007 (2000).

[21] A. Buckley, J. Butterworth, S. Gieseke, D. Grellscheid, S. Höche, H. Hoeth, F. Krauss, L. Lönnblad, E. Nurse, P. Richardson, S. Schumann, M. H. Seymour, T. Sjöstrand, P. Skands, and B. Webber, General-purpose event generators for LHC physics, Phys. Rep. **504**, 145 (2011).

[22] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann, and J. Winter, SHERPA 1.α, a proof-of-concept version, J. High Energy Phys. 02 (2004) 056.

[23] T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert, and J. Winter, Event generation with SHERPA 1.1, J. High Energy Phys. 02 (2009) 007.

[24] F. Krauss, R. Kuhn, and G. Soff, AMEGIC++ 1.0: A Matrix Element Generator In C++, J. High Energy Phys. 02 (2002) 044.

[25] T. Gleisberg and S. Höche, Comix, a new matrix element generator, J. High Energy Phys. 12 (2008) 039.

[26] C. F. Berger, Z. Bern, L. J. Dixon, F. F. Cordero, D. Forde, H. Ita, D. A. Kosower, and D. Maître, Automated implementation of on-shell methods for one-loop amplitudes, Phys. Rev. D **78**, 036003 (2008).
[27] G. Cullen, H. van Deurzen, N. Greiner, G. Heinrich, G. Luisoni, P. Mastrolia, E. Mirabella, G. Ossola, T. Peraro, J. Schlenk, J. F. von Soden-Fraunhofen, and F. Tramontano, G*OSAM*-2.0: a tool for automated one-loop calculations within the Standard Model and beyond, Eur. Phys. J. C **74**, 3001 (2014).
[28] S. Badger, B. Biedermann, P. Uwer, and V. Yundin, Numerical evaluation of virtual corrections to multi-jet production in massless QCD, Comput. Phys. Commun. **184**, 1981 (2013).
[29] F. Cascioli, P. Maierhöfer, and S. Pozzorini, Scattering Amplitudes with Open Loops, Phys. Rev. Lett. **108**, 111601 (2012).
[30] T. Binoth *et al.*, A proposal for a standard interface between Monte Carlo tools and one-loop programs, Comput. Phys. Commun. **181**, 1612 (2010).
[31] The OpenLoops one-loop generator by F. Cascioli, J. Lindert, P. Maierhöfer, and S. Pozzorini is publicly available at http://openloops.hepforge.org.
[32] A. Denner, S. Dittmaier, and L. Hofer, Collier: A Fortran-based complex one-loop library in extended regularizations, Comput. Phys. Commun. **212**, 220 (2017).
[33] A. Denner, S. Dittmaier, and L. Hofer, COLLIER - A fortran-library for one-loop integrals, *Proc. Sci.*, LL2014 (2014) 071.
[34] S. Catani and M. H. Seymour, A general algorithm for calculating jet cross sections in NLO QCD, Nucl. Phys. **B485**, 291 (1997).
[35] S. Catani, S. Dittmaier, M. H. Seymour, and Z. Trocsanyi, The dipole formalism for next–to–leading order QCD calculations with massive partons, Nucl. Phys. **B627**, 189 (2002).
[36] T. Gleisberg and F. Krauss, Automating dipole subtraction for QCD NLO calculations, Eur. Phys. J. C **53**, 501 (2008).
[37] S. Schumann and F. Krauss, A parton shower algorithm based on Catani-Seymour dipole factorisation, J. High Energy Phys. 03 (2008) 038.
[38] S. Höche, S. Schumann, and F. Siegert, Hard photon production and matrix-element parton-shower merging, Phys. Rev. D **81**, 034026 (2010).
[39] Z. Nagy and D. E. Soper, A new parton shower algorithm: Shower evolution, matching at leading and next-to-leading order level, arXiv:hep-ph/0601021.
[40] S. Frixione and B. R. Webber, Matching NLO QCD computations and parton shower simulations, J. High Energy Phys. 06 (2002) 029.
[41] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, A critical appraisal of NLO + PS matching methods, J. High Energy Phys. 09 (2012) 049.
[42] S. Höche, F. Krauss, S. Schumann, and F. Siegert, QCD matrix elements and truncated showers, J. High Energy Phys. 05 (2009) 053.
[43] T. Gehrmann, S. Höche, F. Krauss, M. Schönherr, and F. Siegert, NLO QCD matrix elements + parton showers in $e^+e^- \to$ hadrons, J. High Energy Phys. 01 (2013) 144.
[44] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, QCD matrix elements + parton showers: The NLO case, J. High Energy Phys. 04 (2013) 027.
[45] R. D. Ball, V. Bertone, S. Carrazza, C. S. Deans, L. Del Debbio, S. Forte, A. Guffanti, N. P. Hartland, J. I. Latorre, J. Rojo, and M. Ubiali (NNPDF Collaboration), Parton distributions for the LHC Run II, J. High Energy Phys. 04 (2015) 040.
[46] S. Höche, F. Krauss, M. Schönherr, and F. Siegert, Next-to-leading order matrix elements and truncated showers, arXiv:1009.1477.
[47] E. Bothmann, M. Schönherr, and S. Schumann, Reweighting QCD matrix-element and parton-shower calculations, Eur. Phys. J. C **76**, 590 (2016).
[48] G. Aad *et al.* (ATLAS Collaboration), Measurement of the production cross section of jets in association with a Z boson in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector, J. High Energy Phys. 07 (2013) 032.
[49] V. Khachatryan *et al.* (CMS Collaboration), Measurements of jet multiplicity and differential production cross sections of Z + jets events in proton-proton collisions at $\sqrt{s} = 7$ TeV, Phys. Rev. D **91**, 052008 (2015).
[50] V. Khachatryan *et al.* (CMS Collaboration), Measurements of the differential production cross sections for a Z boson in association with jets in pp collisions at $\sqrt{s} = 8$ TeV, arXiv:1611.03844.
[51] G. Aad *et al.* (ATLAS Collaboration), Measurement of differential production cross-sections for a Z boson in association with b-jets in 7 TeV proton-proton collisions with the ATLAS detector, J. High Energy Phys. 10 (2014) 141.
[52] S. Chatrchyan *et al.* (CMS Collaboration), Measurement of the cross section and angular correlations for associated production of a Z boson with b hadrons in pp collisions at $\sqrt{s} = 7$ TeV, J. High Energy Phys. 12 (2013) 039.
[53] V. Khachatryan *et al.* (CMS Collaboration), Measurements of the associated production of a Z boson and b jets in pp collisions at $\sqrt{s} = 8$ TeV, arXiv:1611.06507.
[54] M. Cacciari, G. P. Salam, and G. Soyez, The anti-k(t) jet clustering algorithm, J. High Energy Phys. 04 (2008) 063.
[55] A. Buckley, J. Butterworth, L. Lönnblad, D. Grellscheid, H. Hoeth, J. Monk, H. Schulz, and F. Siegert, Rivet user manual, Comput. Phys. Commun. **184**, 2803 (2013).
[56] M. Cacciari, G. P. Salam, and G. Soyez, FastJet user manual, Eur. Phys. J. C **72**, 1896 (2012).
[57] S. Dittmaier, M. Kramer, and M. Spira, Higgs radiation off bottom quarks at the Tevatron and the CERN LHC, Phys. Rev. D **70**, 074010 (2004).
[58] S. Dawson, C. B. Jackson, L. Reina, and D. Wackeroth, Exclusive Higgs boson production with bottom quarks at hadron colliders, Phys. Rev. D **69**, 074027 (2004).
[59] S. Dittmaier, A general approach to photon radiation off fermions, Nucl. Phys. **B565**, 69 (2000).